09/PCT/AU99/00220

Patent Office
Canberra

AU99/220

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES, hereby certify that the annexed is a true copy of the Provisional specification in connection with Application No. PP 2634 for a patent by THE WALTER AND ELIZA HALL INSTITUTE OF MEDICAL RESEARCH filed on 27 March 1998.

WITNESS my hand this Twenty-eighth day of April 1999

KIM MARSHALL
MANAGER EXAMINATION SUPPORT AND SALES

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Regulation 3.2

The Walter and Eliza Hall Institute
of Medical Research

A U S T R A L I A

Patents Act 1990

PROVISIONAL SPECIFICATION

for the invention entitled:

"NovelTherapeutic Molecules and Uses Therefor"

The invention is described in the following statement:

NOVEL THERAPEUTIC MOLECULES AND USES THEREFOR

FIELD OF THE INVENTION

The present invention relates to novel molecules expressed during muscle development and to genetic sequences encoding same. More particularly, the present invention relates to novel molecules capable of, *inter alia*, modulating heart and skeletal muscle cell development and to genetic sequences encoding same. Even more particularly, the present invention provides a novel molecule referred to herein as "Csl" and to genetic sequences encoding same. The molecules of the present invention are useful, for example, in therapy, diagnosis and as a screening tool for therapeutic agents capable of modulating muscle cell development.

BACKGROUND OF THE INVENTION

The genetic basis of skeletal muscle development and differentiation has received much attention over the last two decades, especially with regard to the role of the homeobox genes, the basic helix-loop-helix myogenic factors and members of the myocyte enhancer binding factor family (1).

The genetic basis of heart development has enjoyed less attention but is gradually catching up. This is largely due to work on the role of the NK-2 class of homeobox genes. The drosophila NK-2 class homeobox gene *tinman* is so named because flies without this gene, apart from other defects, lack the muscular dorsal vessel that can be loosely described as a heart (2). The search for mammalian homologues of NK-2 revealed a series of genes amongst which the gene Nkx2.5 was found to be expressed in the developing mouse heart (3). Mouse embryos homozygous for a null mutation of Nkx2.5 (Nkx2.5-/-) form a linear heart tube but fail to undergo the morphogenetic process of heart looping during the early stages of cardiogenesis and the embryos die due to hemodynamic insufficiency (4). The cells of the unlooped heart tube show many characteristics of normal cardiocytes such as the expression of many myofilament genes but the knockouts demonstrate that Nkx2.5 is clearly a major determinant of cardiac development.

Already some genes have been identified that fail to be activated in Nkx2.5-/- hearts including the myosin light chain 2V (MLC2V) (4) and the atrial natriuretic factor ANF which is known to have an Nkx2.5 response element (5). An ankyrin repeat-containing nuclear protein CARP is also down regulated in the Nkx2.5-/- model. MLC2v may be regulated directly by CARP and a binding partner YB-1 the ubiquitous transcription factor (6).

In work leading up to the present invention, the inventors have identified a novel gene, designated herein "*Csl*". In accordance with the present invention, *Csl* mRNA is detectable in the heart from about 8 days post-coitum ("dpc") through to adulthood but is absent from hearts in mice carrying a Nkx2.5-/- mutation. It is also expressed in skeletal muscle. The identification of this new gene permits the rational drug design and further identification of a range of molecules for use in therapy, diagnosis, antibody generation and modulation of muscle cell development. These molecules include the product of the *Csl* gene, Csl. Other molecules contemplated herein may act as either agonists or antagonists of Csl's functions and will be useful *inter alia* in attenuation of muscle frailty in aging, treatment of muscular and myotonic dystrophies, the preventation of cardiomyopathy and the regulation of myogenic differentiation.

SUMMARY OF THE INVENTION

One aspect of the present invention provides a nucleic acid molecule comprising a nucleotide sequence corresponding to a gene which is expressed in heart muscle from about 8 dpc in murine species or its equivalents in other mammalia species such as humans.

Another aspect of the present invention provides a nucleic acid molecule comprising a nucleotide sequence encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:2 or a derivative or homologue or mimetic thereof or having at least about 45% or greater similarity to SEQ ID NO:2 or a derivative or homologue or mimetic thereof.

Yet another aspect of the present invention provides a nucleic acid molecule comprising a nucleotide sequence encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:4 or a derivative or homologue or mimetic thereof or having at least about 45% or greater similarity to at least 20 contiguous amino acids in SEQ ID NO:4 or a derivative or homologue or mimetic thereof.

A further aspect of the present invention provides a nucleic acid molecule comprising a nucleotide sequence encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:5 or a derivative or homologue or mimetic thereof or having at least about 45% or greater similarity to at least 20 contiguous amino acids in SEQ ID NO:5 or a derivative or homologue or mimetic thereof.

Yet another further aspect of the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in SEQ ID NO:1 or a derivative or homologue or mimetic thereof capable of hybridising to SEQ ID NO:1 under low stringency conditions at 42°C.

Still another aspect of the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in SEQ ID NO:3 or a derivative or homologue or mimetic thereof capable of hybridising to SEQ ID NO:3 under low stringency conditions at 42°C.

Still yet another aspect of the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in SEQ ID NO:1 or a derivative or homologue or mimetic thereof capable of hybridising to SEQ ID NO:1 under low stringency conditions at 42°C and which encodes an amino acid sequence corresponding to an amino acid sequence set forth in SEQ ID NO:2 or a sequence having at least about 45% similarity to at least 20 contiguous amino acids in SEQ ID NO:2.

Still yet another further aspect of the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in SEQ ID NO:3 or a derivative or homologue or mimetic thereof capable of hybridising to SEQ ID NO:3 under low stringency conditions at 42°C and which encodes an amino acid sequence corresponding to an amino acid sequence set forth in SEQ ID NO:4 or a sequence having at least about 45% similarity to at least 20 contiguous amino acids in SEQ ID NO:4.

Further, the present invention contemplates a nucleic acid molecule comprising a sequence of nucleotides substantially as set forth in SEQ ID NO:1 or 3.

Still yet another further aspect of the present invention contemplates a genomic nucleic acid molecule or a derivative or homologue or mimetic thereof capable of hybridising to SEQ ID NO:1 or 3 or a derivative or homologue or mimetic thereof under low stringency conditions at 42°C.

In another further aspect the nucleotide sequence corresponding to *Csl* is a cDNA sequence comprising a sequence of nucleotides as set forth in SEQ ID NO:1 or is a derivative or homologue or mimetic thereof including a nucleotide sequence having similarity to SEQ ID NO:1.

In yet another further aspect the nucleotide sequence corresponding to *Csl* is a cDNA sequence comprising a sequence of nucleotides as set forth in SEQ ID NO:3 or is a derivative or homologue or mimetic thereof including a nucleotide sequence having similarity to SEQ ID NO:3.

In still yet another further aspect the nucleotide sequence corresponding to *Csl* is a genomic nucleotide sequence or a derivative, homologue or mimetic thereof corresponding to a gene map as set forth in Figure 1.

In still another further aspect the nucleotide sequence corresponding to *Csl* is a genomic nucleotide sequence or a derivative, homologue or mimetic thereof having exon regions of

which 5 comprise:

Exon 1 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 19498 and ending at residue 19327 of the human cosmid clone U228D4;

Exon 2 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 15687 and ending at residue 15631 of the human cosmid clone U228D4;

Exon 3 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 5220 and ending at residue 5133 of the human cosmid clone U228D4;

Exon 4 comprising the nucleotide sequence beginning at residue 35384 and ending at residue 35236 of the human cosmid clone U112E8; and

Exon 5 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 4101 and ending at residue 3680 of the human cosmid clone U112E8.

The Accession numbers corresponding to human cosmid clones U228D4 and U112E8 are U73509 and U73508, respectively.

Further, the present invention provides a nucleotide sequence corresponding to *Csl* is a genomic nucleotide sequence or a derivative, homologue or mimetic thereof corresponding to a gene map as set forth in Figure 2.

Another aspect of the present invention is directed to a isolated nucleic acid molecule encoding *Csl* or a derivative, homologue or mimetic thereof, said nucleic acid molecule selected from the list consisting of:

(i) A nucleic acid molecule comprising a nucleotide sequence encoding the amino acid sequence set forth in one or more of SEQ ID NO:2 or 4 or 5 or a derivative or homologue or mimetic thereof or having at least about 45% similarity to at least 20 contiguous amino acids in one or more of SEQ ID NO: 2 or 4 or 5.

(ii) A nucleic acid molecule comprising a nucleotide sequence substantially as set forth in one or more of SEQ ID NO:1 or 3 or a derivative or homologue or mimetic thereof.

(iii) A nucleic acid molecule comprising a nucleotide sequence corresponding to a gene map as set forth in Figures 1 or 2 or a derivative or homologue thereof.

(iv) A nucleic acid molecule comprising 5 exons wherein:

Exon 1 comprises a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 19498 and ending at residue 19327 of the human cosmid clone U228D4;

Exon 2 comprises a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 15687 and ending at residue 15631 of the human cosmid clone U228D4;

Exon 3 comprises a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 5220 and ending at residue 5133 of the human cosmid clone U228D4;

Exon 4 comprises the nucleotide sequence beginning at residue 35384 and ending at residue 35236 of the human cosmid clone U112E8; and

Exon 5 comprises a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 4101 and ending at residue 3680 of the human cosmid clone U112E8;

or a derivative, homologue or mimetic thereof.

(v) A nucleic acid molecule capable of hybridising under low stringency conditions at 42°C to the nucleotide sequence substantially as set forth in one or more of SEQ ID NO:1 or 3 or a derivative or homologue or mimetic thereof and encoding an amino acid sequence corresponding to an amino acid sequence as set forth in one or more of SEQ ID NO:2 or 4, respectively, or a derivative or homologue or mimetic thereof or a sequence having at least about 45% similarity to at least 20 contiguous amino acids in one or more of SEQ ID NO:2 or 4, respectively.

(vi) A nucleic acid molecule capable of hybridising to the nucleic acid molecule of paragraphs (i), (ii) or (v) under low stringency conditions at 42°C and encoding an amino acid sequence having at least about 45% similarity to at least 20 contiguous amino acids in one or more of SEQ ID NO:2 or 4.

(vii) A nucleic acid molecule capable of hybridising to the nucleic acid molecule of paragraphs (iii) or (iv) under low stringency conditions at 42°C.

(viii) A derivative or homologue of the nucleic acid molecule of paragraphs (i) or (ii) or (iii) or (iv) or (v) or (vi) or (vii).

Yet another aspect of the present invention is directed to an isolated protein selected from the list consisting of:

(i) A protein having an amino acid sequence substantially as set forth in one or more of SEQ ID NO:2 or 4 or 5 or derivative or homologue or mimetic thereof or a sequence having at least about 45% similarity to at least 20 contiguous amino acids in one or more of SEQ ID NO:2 or 4 or 5.

(ii) A protein encoded by a nucleotide sequence substantially as set forth in one or more of SEQ ID NO:1 or 3 or a derivative or homologue or mimetic thereof or a sequence encoding an amino acid sequence having at least about 45% similarity to at least 20 contiguous amino acids in one or more of SEQ ID NO:2 or 4, respectively.

(iii) A protein encoded by a nucleotide sequence corresponding to a gene map as set forth in Figure 1 or 2 or a derivative or homologue or mimetic thereof.

(iv) A protein encoded by a nucleotide sequence comprising 5 exons wherein:

Exon 1 comprises a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 19498 and ending at residue 19327 of the human cosmid clone U228D4;

Exon 2 comprises a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 15687 and ending at residue 15631 of the human cosmid clone U228D4;

Exon 3 comprises a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 5220 and ending at residue 5133 of the human cosmid clone U228D4;

Exon 4 comprises the nucleotide sequence beginning at residue 35384 and ending at residue 35236 of the human cosmid clone U112E8; and

Exon 5 comprises a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 4101 and ending at residue 3680 of the human cosmid clone U112E8;

or a derivative, homologue or mimetic thereof.

(v) A protein encoded by a nucleic acid molecule capable of hybridising to the nucleotide sequence as set forth in one of SEQ ID NO:1 or 3 or a derivative or homologue or mimetic thereof under low stringency conditions at 42°C and which encodes an amino acid sequence substantially as set forth in SEQ ID NO:2 or 4, respectively, or a derivative or homologue or mimetic thereof or an amino acid sequence having at least about 45% similarity to at least 20 contiguous amino acids in one or more SEQ ID NO:2 or 4, respectively.

(vi) A protein as defined in paragraphs (i) or (ii) or (iii) or (iv) or (v) in a homodimeric form.

(vii) A protein as defined in paragraphs (i) or (ii) or (iii) or (iv) or (v) in heterodimeric form.

Yet another aspect of the present invention contemplates a method for modulating activity of Csl in a mammal, said method comprising administering to said mammal a modulating effective amount of an agent for a time and under conditions sufficient to increase or decrease Csl activity.

A further aspect of the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of an agent capable of modulating the expression of *Csl* or a derivative or homologue or mimetic thereof for a time and under conditions sufficient to modulate muscle cell development.

In another aspect the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of an agent capable of modulating the activity of Csl or a derivative, homologue or mimetic thereof for a time and under conditions sufficient to modulate muscle cell development.

In another aspect the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of Csl or a derivative, homologue or mimetic thereof for a time and under conditions sufficient to modulate muscle cell development.

Yet another aspect the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of *Csl* or a derivative, homologue or mimetic thereof for a time and under conditions sufficient to modulate muscle cell development.

In yet another aspect the present invention relates to the use of an agent capable of modulating the expression of *Csl* or a derivative, homologue or mimetic thereof in the manufacture of a medicament for the modulation of muscle cell development.

Another aspect of the present invention relates to the use of an agent capable of modulating the expression of Csl or a derivative, homologue or mimetic thereof in the manufacture of a medicament for the modulation of muscle cell development.

A further aspect of the present invention relates to the use of Csl or *Csl* in the manufacture of a medicament for the modulation of muscle cell development.

Still yet another aspect of the present invention relates to agents for use in modulating *Csl* or a derivative, homologue or mimetic thereof expression wherein modulating expression of said *Csl* modulates muscle cell development.

A further aspect of the present invention relates to agents for use in modulating Csl or a derivative, homologue or mimetic thereof expression wherein modulating expression of said Csl modulates muscle cell development.

Another aspect of the present invention relates to Csl or *Csl* for use in modulating muscle cell development.

In yet another further aspect the present invention contemplates a pharamceutical composition *Csl*, Csl or an agent capable of modulating *Csl* expression or Csl activity together with one or more pharamaceutically acceptable carriers and/or diluents. *Csl*, Csl or said agent are referred to as the active ingredients.

Bibliographic details of the publications numerically referred to in this specification are collected at the end of the description. Sequence Identity Numbers (SEQ ID NOs.) for the nucleotide and amino acid sequences referred to in the specification are defined following the Examples. A summary of the sequences with given SEQ ID NOs is provided before the Examples.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

DETAILED DESCRIPTION OF THE INVENTION

Accordingly, one aspect of the present invention provides a nucleic acid molecule comprising a nucleotide sequence corresponding to a gene which is expressed in heart muscle from about

8 dpc in murine species or its equivalents in other mammalia species such as humans.

More particularly, the present invention provides a nucleic acid molecule comprising a nucleotide sequence encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:2 or a derivative or homologue or mimetic thereof or having at least about 45% or greater similarity to SEQ ID NO:2 or a derivative or homologue or mimetic thereof.

Even more particularly the present invention provides a nucleic acid molecule comprising a nucleotide sequence encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:4 or a derivative or homologue or mimetic thereof or having at least about 45% or greater similarity to at least 20 contiguous amino acids in SEQ ID NO:4 or a derivative or homologue or mimetic thereof.

Yet even more particularly the present invention provides a nucleic acid molecule comprising a nucleotide sequence encoding or complementary to a sequence encoding an amino acid sequence substantially as set forth in SEQ ID NO:5 or a derivative or homologue or mimetic thereof or having at least about 45% or greater similarity to at least 20 contiguous amino acids in SEQ ID NO:5 or a derivative or homologue or mimetic thereof.

The term "similarity" as used herein includes exact identity between compared sequences at the nucleotide or amino acid level. Where there is non-identity at the nucleotide level, "similarity" includes differences between sequences which result in different amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. Where there is non-identity at the amino acid level, "similarity" includes amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels.

Another aspect of the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in SEQ ID NO:1 or a derivative or homologue or mimetic thereof capable of hybridising to SEQ ID NO:1 under low stringency conditions

at 42°C.

Still another aspect of the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in SEQ ID NO:3 or a derivative or homologue or mimetic thereof capable of hybridising to SEQ ID NO:3 under low stringency conditions at 42°C.

Reference herein to a low stringency at 42°C includes and encompasses from at least about 1% v/v to at least about 15% v/v formamide and from at least about 1M to at least about 2M salt for hybridisation, and at least about 1M to at least about 2M salt for washing conditions. Alternative stringency conditions may be applied where necessary, such as medium stringency, which includes and encompasses from at least about 16% v/v to at least about 30% v/v formamide and from at least about 0.5M to at least about 0.9M salt for hybridisation, and at least about 0.5M to at least about 0.9M salt for washing conditions, or high stringency, which includes and encompasses from at least about 31% v/v to at least about 50% v/v formamide and from at least about 0.01M to at least about 0.15M salt for hybridisation, and at least about 0.01M to at least about 0.15M salt for washing conditions.

Preferably the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in SEQ ID NO:1 or a derivative or homologue or mimetic thereof capable of hybridising to SEQ ID NO:1 under low stringency conditions at 42°C and which encodes an amino acid sequence corresponding to an amino acid sequence set forth in SEQ ID NO:2 or a sequence having at least about 45% similarity to at least 20 contiguous amino acids in SEQ ID NO:2.

Preferably the present invention contemplates a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in SEQ ID NO:3 or a derivative or homologue or mimetic thereof capable of hybridising to SEQ ID NO:3 under low stringency conditions at 42°C and which encodes an amino acid sequence corresponding to an amino acid sequence set forth in SEQ ID NO:4 or a sequence having at least about 45% similarity to at least 20 contiguous amino acids in SEQ ID NO:4.

More particularly, the present invention contemplates a nucleic acid molecule comprising a sequence of nucleotides substantially as set forth in SEQ ID NO:1 or 3.

The nucleic acid molecule according to this aspect of the present invention corresponds herein to "*Csl*". This gene has been determined, in accordance with the present invention, to undergo expression in heart muscle from 7.5 dpc in normal mice but not in the hearts of mutants which lack the major determinant of cardiac development - Nkx2.5. It is thereby thought to form a novel member of the group of genes involved in muscle formation and in particular heart formation. The product of the *Csl* gene is referred to herein as Csl. Murine Csl is defined by the amino acid sequence set forth in SEQ ID NO:2, human Csl is defined by the amino acid sequence set forth in SEQ ID NO:4 and Xenopus Csl is defined by the amino acid sequence set forth in SEQ ID NO:5. The cDNA nucleotide sequence for murine and human Csl are defined by the nucleotide sequences set forth in SEQ ID NO:1 and SEQ ID NO:3, respectively.

The nucleic acid molecule encoding Csl is preferably a sequence of deoxyribonucleic acids such as a cDNA sequence or a genomic sequence. A genomic sequence may also comprise exons and introns. A genomic sequence may also include a promoter region or other regulatory regions.

Another aspect of the present invention contemplates a genomic nucleic acid molecule or a derivative or homologue or mimetic thereof capable of hybridising to SEQ ID NO:1 or 3 or a derivative or homologue or mimetic thereof under low stringency conditions at 42°C.

Reference herein to Csl and *Csl* should be understood as a reference to all forms of Csl and *Csl*,respectively, including, for example, any peptide and cDNA isoforms which arise from alternative splicing of *Csl* mRNA. Reference hereinafter to Csl and *Csl* includes reference to derivatives thereof.

The term "protein" should be understood to encompass peptides polypeptides and proteins. The protein may be glycosylated or unglycosylated and/or may contain a range of other

molecules fused, linked, bound or otherwise associated to the protein such as amino acids, lipids, carbohydrates or other peptides, polypeptides or proteins. Reference hereinafter to a "protein" includes a protein comprising a sequence of amino acids as well as a protein associated with other molecules such as amino acids, lipids, carbohydrates or other peptides, polypeptides or proteins.

The protein and/or gene is preferably from a human, primate, livestock animal (eg. sheep, pig, cow, horse, donkey) laboratory test animal (eg. mouse, rat, rabbit, guinea pig) companion animal (eg. dog, cat), captive wild animal (eg. fox, kangaroo, deer), aves (eg. chicken, geese, duck, emu, ostrich), reptile or fish.

Derivatives include fragments, parts, portions, chemical equivalents, mutants, homologs, mimetics from natural, synthetic or recombinant sources including fusion proteins. Derivatives may be derived from insertion, deletion or substitution of amino acids. Amino acid insertional derivatives include amino and/or carboxylic terminal fusions as well as intrasequence insertions of single or multiple amino acids. Insertional amino acid sequence variants are those in which one or more amino acid residues are introduced into a predetermined site in the protein although random insertion is also possible with suitable screening of the resulting product. Deletional variants are characterized by the removal of one or more amino acids from the sequence. Substitutional amino acid variants are those in which at least one residue in the sequence has been removed and a different residue inserted in its place. Additions to amino acid sequences including fusions with other peptides, polypeptides or proteins.

The derivatives of Csl include fragments having particular epitopes of parts of the entire Csl protein fused to peptides, polypeptides or other proteinaceous or non-proteinaceous molecules. For example, Csl or derivative thereof may be fused to a molecule to facilitate its entry into a cell. Analogs of Csl contemplated herein include, but are not limited to, modification to side chains, incorporating of unnatural amino acids and/or their derivatives during peptide, polypeptide or protein synthesis and the use of crosslinkers and other methods which impose conformational constraints on the proteinaceous molecules or their

analogs. Derivatives of nucleic acid sequences may similarly be derived from single or multiple nucleotide substitutions, deletions and/or additions including fusion with other nucleic acid molecules. The derivatives of the nucleic acid molecules of the present invention include oligonucleotides, PCR primers, antisense molecules, molecules suitable for use in cosuppression and fusion of nucleic acid molecules.

Examples of side chain modifications contemplated by the present invention include modifications of amino groups such as by reductive alkylation by reaction with an aldehyde followed by reduction with $NaBH_4$; amidination with methylacetimidate; acylation with acetic anhydride; carbamoylation of amino groups with cyanate; trinitrobenzylation of amino groups with 2, 4, 6-trinitrobenzene sulphonic acid (TNBS); acylation of amino groups with succinic anhydride and tetrahydrophthalic anhydride; and pyridoxylation of lysine with pyridoxal-5-phosphate followed by reduction with $NaBH_4$.

The guanidine group of arginine residues may be modified by the formation of heterocyclic condensation products with reagents such as 2,3-butanedione, phenylglyoxal and glyoxal.

The carboxyl group may be modified by carbodiimide activation *via* O-acylisourea formation followed by subsequent derivitisation, for example, to a corresponding amide.

Sulphydryl groups may be modified by methods such as carboxymethylation with iodoacetic acid or iodoacetamide; performic acid oxidation to cysteic acid; formation of a mixed disulphides with other thiol compounds; reaction with maleimide, maleic anhydride or other substituted maleimide; formation of mercurial derivatives using 4-chloromercuribenzoate, 4-chloromercuriphenylsulphonic acid, phenylmercury chloride, 2-chloromercuri-4-nitrophenol and other mercurials; carbamoylation with cyanate at alkaline pH.

Tryptophan residues may be modified by, for example, oxidation with N-bromosuccinimide or alkylation of the indole ring with 2-hydroxy-5-nitrobenzyl bromide or sulphenyl halides. Tyrosine residues on the other hand, may be altered by nitration with tetranitromethane to form a 3-nitrotyrosine derivative.

Modification of the imidazole ring of a histidine residue may be accomplished by alkylation with iodoacetic acid derivatives or N-carboethoxylation with diethylpyrocarbonate.

Examples of incorporating unnatural amino acids and derivatives during protein synthesis include, but are not limited to, use of norleucine, 4-amino butyric acid, 4-amino-3-hydroxy-5-phenylpentanoic acid, 6-aminohexanoic acid, t-butylglycine, norvaline, phenylglycine, ornithine, sarcosine, 4-amino-3-hydroxy-6-methylheptanoic acid, 2-thienyl alanine and/or D-isomers of amino acids. A list of unnatural amino acid contemplated herein is shown in Table 1.

TABLE 1

Non-conventional amino acid	Code	Non-conventional amino acid	Code
α-aminobutyric acid	Abu	L-N-methylalanine	Nmala
α-amino-α-methylbutyrate	Mgabu	L-N-methylarginine	Nmarg
aminocyclopropane-carboxylate	Cpro	L-N-methylasparagine	Nmasn
		L-N-methylaspartic acid	Nmasp
aminoisobutyric acid	Aib	L-N-methylcysteine	Nmcys
aminonorbornyl-carboxylate	Norb	L-N-methylglutamine	Nmgln
		L-N-methylglutamic acid	Nmglu
cyclohexylalanine		Chexa L-N-methylhistidine	Nmhis
cyclopentylalanine	Cpen	L-N-methylisolleucine	Nmile
D-alanine	Dal	L-N-methylleucine	Nmleu
D-arginine	Darg	L-N-methyllysine	Nmlys
D-aspartic acid	Dasp	L-N-methylmethionine	Nmmet
D-cysteine	Dcys	L-N-methylnorleucine	Nmnle
D-glutamine	Dgln	L-N-methylnorvaline	Nmnva
D-glutamic acid	Dglu	L-N-methylornithine	Nmorn
D-histidine	Dhis	L-N-methylphenylalanine	Nmphe
D-isoleucine	Dile	L-N-methylproline	Nmpro
D-leucine	Dleu	L-N-methylserine	Nmser
D-lysine	Dlys	L-N-methylthreonine	Nmthr
D-methionine	Dmet	L-N-methyltryptophan	Nmtrp
D-ornithine	Dorn	L-N-methyltyrosine	Nmtyr
D-phenylalanine	Dphe	L-N-methylvaline	Nmval
D-proline	Dpro	L-N-methylethylglycine	Nmetg
D-serine	Dser	L-N-methyl-t-butylglycine	Nmtbug
D-threonine	Dthr	L-norleucine	Nle

	D-tryptophan	Dtrp	L-norvaline	Nva
	D-tyrosine	Dtyr	α-methyl-aminoisobutyrate	Maib
	D-valine	Dval	α-methyl-γ-aminobutyrate	Mgabu
	D-α-methylalanine	Dmala	α-methylcyclohexylalanine	Mchexa
5	D-α-methylarginine	Dmarg	α-methylcylcopentylalanine	Mcpen
	D-α-methylasparagine	Dmasn	α-methyl-α-napthylalanine	Manap
	D-α-methylaspartate	Dmasp	α-methylpenicillamine	Mpen
	D-α-methylcysteine	Dmcys	N-(4-aminobutyl)glycine	Nglu
	D-α-methylglutamine	Dmgln	N-(2-aminoethyl)glycine	Naeg
10	D-α-methylhistidine	Dmhis	N-(3-aminopropyl)glycine	Norn
	D-α-methylisoleucine	Dmile	N-amino-α-methylbutyrate	Nmaabu
	D-α-methylleucine	Dmleu	α-napthylalanine	Anap
	D-α-methyllysine	Dmlys	N-benzylglycine	Nphe
	D-α-methylmethionine	Dmmet	N-(2-carbamylethyl)glycine	Ngln
15	D-α-methylornithine	Dmorn	N-(carbamylmethyl)glycine	Nasn
	D-α-methylphenylalanine	Dmphe	N-(2-carboxyethyl)glycine	Nglu
	D-α-methylproline	Dmpro	N-(carboxymethyl)glycine	Nasp
	D-α-methylserine	Dmser	N-cyclobutylglycine	Ncbut
	D-α-methylthreonine	Dmthr	N-cycloheptylglycine	Nchep
20	D-α-methyltryptophan	Dmtrp	N-cyclohexylglycine	Nchex
	D-α-methyltyrosine	Dmty	N-cyclodecylglycine	Ncdec
	D-α-methylvaline	Dmval	N-cylcododecylglycine	Ncdod
	D-N-methylalanine	Dnmala	N-cyclooctylglycine	Ncoct
	D-N-methylarginine	Dnmarg	N-cyclopropylglycine	Ncpro
25	D-N-methylasparagine	Dnmasn	N-cycloundecylglycine	Ncund
	D-N-methylaspartate	Dnmasp	N-(2,2-diphenylethyl)glycine	Nbhm
	D-N-methylcysteine	Dnmcys	N-(3,3-diphenylpropyl)glycine	Nbhe
	D-N-methylglutamine	Dnmgln	N-(3-guanidinopropyl)glycine	Narg
	D-N-methylglutamate	Dnmglu	N-(1-hydroxyethyl)glycine	Nthr
30	D-N-methylhistidine	Dnmhis	N-(hydroxyethyl))glycine	Nser

D-N-methylisoleucine	Dnmile	N-(imidazolylethyl))glycine	Nhis
D-N-methylleucine	Dnmleu	N-(3-indolylyethyl)glycine	Nhtrp
D-N-methyllysine	Dnmlys	N-methyl-γ-aminobutyrate	Nmgabu
N-methylcyclohexylalanine	Nmchexa	D-N-methylmethionine	Dnmmet
D-N-methylornithine	Dnmorn	N-methylcyclopentylalanine	Nmcpen
N-methylglycine	Nala	D-N-methylphenylalanine	Dnmphe
N-methylaminoisobutyrate	Nmaib	D-N-methylproline	Dnmpro
N-(1-methylpropyl)glycine	Nile	D-N-methylserine	Dnmser
N-(2-methylpropyl)glycine	Nleu	D-N-methylthreonine	Dnmthr
D-N-methyltryptophan	Dnmtrp	N-(1-methylethyl)glycine	Nval
D-N-methyltyrosine	Dnmtyr	N-methyla-napthylalanine	Nmanap
D-N-methylvaline	Dnmval	N-methylpenicillamine	Nmpen
γ-aminobutyric acid	Gabu	N-(*p*-hydroxyphenyl)glycine	Nhtyr
L-*t*-butylglycine	Tbug	N-(thiomethyl)glycine	Ncys
L-ethylglycine	Etg	penicillamine	Pen
L-homophenylalanine	Hphe	L-α-methylalanine	Mala
L-α-methylarginine	Marg	L-α-methylasparagine	Masn
L-α-methylaspartate	Masp	L-α-methyl-*t*-butylglycine	Mtbug
L-α-methylcysteine	Mcys	L-methylethylglycine	Metg
L-α-methylglutamine	Mgln	L-α-methylglutamate	Mglu
L-α-methylhistidine	Mhis	L-α-methylhomophenylalanine	Mhphe
L-α-methylisoleucine	Mile	N-(2-methylthioethyl)glycine	Nmet
L-α-methylleucine	Mleu	L-α-methyllysine	Mlys
L-α-methylmethionine	Mmet	L-α-methylnorleucine	Mnle
L-α-methylnorvaline	Mnva	L-α-methylornithine	Morn
L-α-methylphenylalanine	Mphe	L-α-methylproline	Mpro
L-α-methylserine	Mser	L-α-methylthreonine	Mthr
L-α-methyltryptophan	Mtrp	L-α-methyltyrosine	Mtyr
L-α-methylvaline	Mval	L-N-methylhomophenylalanine	Nmhphe
N-(N-(2,2-diphenylethyl)	Nnbhm	N-(N-(3,3-diphenylpropyl)	Nnbhe

carbamylmethyl)glycine 1-carboxy-1-(2,2-diphenyl-Nmbc ethylamino)cyclopropane	carbamylmethyl)glycine

Crosslinkers can be used, for example, to stabilise 3D conformations, using homo-bifunctional crosslinkers such as the bifunctional imido esters having $(CH_2)_n$ spacer groups with n=1 to n=6, glutaraldehyde, N-hydroxysuccinimide esters and hetero-bifunctional reagents which usually contain an amino-reactive moiety such as N-hydroxysuccinimide and another group specific-reactive moiety such as maleimido or dithio moiety (SH) or carbodiimide (COOH). In addition, peptides can be conformationally constrained by, for example, incorporation of C_α and N_α-methylamino acids, introduction of double bonds between C_α and C_β atoms of amino acids and the formation of cyclic peptides or analogues by introducing covalent bonds such as forming an amide bond between the N and C termini, between two side chains or between a side chain and the N or C terminus.

The nucleic acid molecule of the present invention is preferably in isolated form or ligated to a vector, such as an expression vector. By "isolated" is meant a nucleic acid molecule having undergone at least one purification step and this is conveniently defined, for example, by a composition comprising at least about 10% subject nucleic acid molecule, preferably at least about 20%, more preferably at least about 30%, still more preferably at least about 40-50%, even still more preferably at least about 60-70%, yet even still more preferably 80-90% or greater of subject nucleic acid molecule relative to other components as determined by molecular weight, encoding activity, nucleotide sequence, base composition or other convenient means. The nucleic acid molecule of the present invention may also be considered, in a preferred embodiment, to be biologically pure.

In a particularly preferred embodiment the nucleotide sequence corresponding to *Csl* is a cDNA sequence comprising a sequence of nucleotides as set forth in SEQ ID NO:1 or is a derivative or homologue or mimetic thereof including a nucleotide sequence having similarity to SEQ ID NO:1.

In another particularly preferred embodiment, the nucleotide sequence corresponding to *Csl* is a cDNA sequence comprising a sequence of nucleotides as set forth in SEQ ID NO:3 or is a derivative or homologue or mimetic thereof including a nucleotide sequence having similarity to SEQ ID NO:3.

In yet another particularly preferred embodiment the nucleotide sequence corresponding to *Csl* is a genomic nucleotide sequence or a derivative, homologue or mimetic thereof corresponding to a gene map as set forth in Figure 1.

In still yet another particularly preferred embodiment the nucleotide sequence corresponding to *Csl* is a genomic nucleotide sequence or a derivative, homologue or mimetic thereof having exon regions of which 5 comprise:

Exon 1 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 19498 and ending at residue 19327 of the human cosmid clone U228D4;

Exon 2 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 15687 and ending at residue 15631 of the human cosmid clone U228D4;

Exon 3 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 5220 and ending at residue 5133 of the human cosmid clone U228D4;

Exon 4 comprising the nucleotide sequence beginning at residue 35384 and ending at residue 35236 of the human cosmid clone U112E8; and

Exon 5 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 4101 and ending at residue 3680 of the human cosmid clone U112E8.

The Accession numbers corresponding to human cosmid clones U228D4 and U112E8 are U73509 and U73508, respectively.

In another particularly preferred embodiment the nucleotide sequence corresponding to *Csl* is a genomic nucleotide sequence or a derivative, homologue or mimetic thereof corresponding to a gene map as set forth in Figure 2.

A derivative of a nucleic acid molecule of the present invention also includes a nucleic acid molecule capable of hybridising to a nucleotide sequences set forth in one of SEQ ID NO:1 or 3 or to a nucleotide sequence corresponding to the gene maps set out in Figures 1 or 2 or to a nucleotide sequence comprising 5 exons as hereinbefore defined under low stringency conditions. Preferably, low stringency is at 42°C.

In another embodiment the present invention is directed to a isolated nucleic acid molecule encoding *Csl* or a derivative, homologue or mimetic thereof, said nucleic acid molecule selected from the list consisting of:

(i) A nucleic acid molecule comprising a nucleotide sequence encoding the amino acid sequence set forth in one or more of SEQ ID NO:2 or 4 or 5 or a derivative or homologue or mimetic thereof or having at least about 45% similarity to at least 20 contiguous amino acids in one or more of SEQ ID NO: 2 or 4 or 5.

(ii) A nucleic acid molecule comprising a nucleotide sequence substantially as set forth in one or more of SEQ ID NO:1 or 3 or a derivative or homologue or mimetic thereof.

(iii) A nucleic acid molecule comprising a nucleotide sequence corresponding to a gene map as set forth in Figures 1 or 2 or a derivative or homologue or mimetic thereof.

(iv) A nucleic acid molecule having exon regions of which 5 comprise:

Exon 1 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 19498 and ending at residue 19327 of the human cosmid clone U228D4;

Exon 2 comprising a nucleotide sequence corresponding to the nucleotide

sequence beginning at residue 15687 and ending at residue 15631 of the human cosmid clone U228D4;

Exon 3 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 5220 and ending at residue 5133 of the human cosmid clone U228D4;

Exon 4 comprising the nucleotide sequence beginning at residue 35384 and ending at residue 35236 of the human cosmid clone U112E8; and

Exon 5 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 4101 and ending at residue 3680 of the human cosmid clone U112E8;

or a derivative, homologue or mimetic thereof.

(v) A nucleic acid molecule capable of hybridising under low stringency conditions at 42°C to the nucleotide sequence substantially as set forth in one or more of SEQ ID NO:1 or 3 or a derivative or homologue or mimetic thereof and encoding an amino acid sequence corresponding to an amino acid sequence as set forth in one or more of SEQ ID NO:2 or 4, respectively, or a derivative or homologue or mimetic thereof or a sequence having at least about 45% similarity to at least 20 contiguous amino acids in one or more of SEQ ID NO:2 or 4, respectively.

(vi) A nucleic acid molecule capable of hybridising to the nucleic acid molecule of paragraphs (i), (ii) or (v) under low stringency conditions at 42°C and encoding an amino acid sequence having at least about 45% similarity to at least 20 contiguous amino acids in one or more of SEQ ID NO:2 or 4.

(vii) A nucleic acid molecule capable of hybridising to the nucleic acid molecule of paragraphs (iii) or (iv) under low stringency conditions at 42°C.

(viii) A derivative or homologue of the nucleic acid molecule of paragraphs (i) or (ii) or (iii) or (iv) or (v) or (vi) or (vii).

The nucleic acid molecule may be ligated to an expression vector capable of expression in a prokaryotic cell (e.g. *E.coli*) or a eukaryotic cell (e.g. yeast cells, fungal cells, insect cells, mammalian cells or plant cells). The nucleic acid molecule may be ligated or fused or otherwise associated with a nucleic acid molecule encoding another entity such as, for example, a signal peptide or a cytokine.

The present invention extends to the expression product of the nucleic acid molecule hereinbefore defined.

The expression product is Csl having an amino acid sequence set forth in one or more of SEQ ID NO:2 or 4 or 5 or is a derivative or homologue or mimetic thereof as defined above or is a derivative, homologue or mimetic having an amino acid sequence of at least about 45% similarity to at least 20 contiguous amino acids in the amino acid sequence as set forth in one or more of SEQ ID NO:2 or 4 or 5 or a derivative or homologue or mimetic thereof.

Another aspect of the present invention is directed to an isolated protein selected from the list consisting of:

(i) A protein having an amino acid sequence substantially as set forth in one or more of SEQ ID NO:2 or 4 or 5 or derivative or homologue or mimetic thereof or a sequence having at least about 45% similarity to at least 20 contiguous amino acids in one or more of SEQ ID NO:2 or 4 or 5.

(ii) A protein encoded by a nucleotide sequence substantially as set forth in one or more of SEQ ID NO:1 or 3 or a derivative or homologue or mimetic thereof or a sequence encoding an amino acid sequence having at least about 45% similarity to at least 20 contiguous amino acids in one or more of SEQ ID NO:2 or 4, respectively.

(iii) A protein encoded by a nucleotide sequence corresponding to a gene map as set forth in Figure 1 or 2 or a derivative or homologue or mimetic thereof.

(iv) A protein encoded by a nucleotide sequence having exon regions of which 5 comprise:

Exon 1 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 19498 and ending at residue 19327 of the human cosmid clone U228D4;

Exon 2 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 15687 and ending at residue 15631 of the human cosmid clone U228D4;

Exon 3 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 5220 and ending at residue 5133 of the human cosmid clone U228D4;

Exon 4 comprising the nucleotide sequence beginning at residue 35384 and ending at residue 35236 of the human cosmid clone U112E8; and

Exon 5 comprising a nucleotide sequence corresponding to the nucleotide sequence beginning at residue 4101 and ending at residue 3680 of the human cosmid clone U112E8;

or a derivative, homologue or mimetic thereof.

(v) A protein encoded by a nucleic acid molecule capable of hybridising to the nucleotide sequence as set forth in one of SEQ ID NO:1 or 3 or a derivative or homologue or mimetic thereof under low stringency conditions at 42°C and which encodes an amino acid sequence substantially as set forth in SEQ ID NO:2 or 4, respectively, or a derivative or homologue or mimetic thereof or an amino acid sequence having at least about 45% similarity to at least 20 contiguous amino acids in one or more SEQ ID NO:2 or 4, respectively.

(vi) A protein as defined in paragraphs (i) or (ii) or (iii) or (iv) or (v) in a homodimeric form.

(vii) A protein as defined in paragraphs (i) or (ii) or (iii) or (iv) or (v) in heterodimeric form.

The Csl of the present invention may be in multimeric form meaning that two or more molecules are associated together. Where the same Csl molecules are associated together, the complex is a homomultimer. An example of a homomultimer is a homodimer. Where at least one Csl is associated with at least one non-Csl molecule, then the complex is a heteromultimer such as a heterodimer. A heteromultimer may include for example, another molecule capable of modulating muscle cell differentiation.

Although not intending to limit the invention to any one theory or mode of action, Csl expression becomes activated upon muscle differentiation. This protein has been determined to co-localise with muscle structural proteins such as, but not limited to, myosin. Csl is a component of a signalling system that modulates gene expression in the muscles. It can thereby promote myogenic differentiation and is therefore an important regulator of the myogenic program. Csl is thought to respond to the withdrawal of myoblasts from the cell cycle, to interact with elements of the cytoskeleton and to play a role in the differentiation of contractile cells.

Csl is thought to be part of a monitoring system for maintenance of muscle homeostasis, a system which may also function in development to coordinate muscle size and pattern. The myofilament is constantly monitoring itself, it grows larger if used heavily, and is lost if used less. Experimental data indicate that a feedback system, working at both the transcriptional and translational levels, ensures correct homeostasis. First, the level of myofilament protein in the heart remains constant for a particular set of physiological parameters, despite natural or experimental variation in the level of transcription of individual myofilament genes. Second, enforced expression of skeletal α-actin in C2C12 myoblasts causes transcriptional activation of other thin filament genes. Third, mechanical stretch in cardiac and skeletal muscle, one of the key sensors of load, induces a complex series of signal transduction events which leads to gene expression and muscular hypertrophy. Csl is thought to function in these events.

The present invention contemplates, therefore, a method for modulating expression of *Csl* in a mammal, said method comprising contacting the *Csl* gene with an effective amount of an

agent for a time and under conditions sufficient to up-regulate or down-regulate or otherwise modulate expression of *Csl*. For example, *Csl* antisense sequences such as oligonucleotides may be introduced into a cell to down-regulate the functional activity of Csl. Conversely, a nucleic acid molecule encoding Csl or a derivative thereof may be introduced to enhance the functional activity of Csl in any cell expressing the endogenous *Csl* genes.

The term "modulating" means up-regulating or down-regulating. Accordingly, although the preferred method is to induce muscle cell development, the inhibition of muscle cell development may also be desired to reverse maladaptive hypertrophy.

An "effective amount" means an amount necessary to at least partly attain the desired response.

Another aspect of the present invention contemplates a method for modulating activity of Csl in a mammal, said method comprising administering to said mammal a modulating effective amount of an agent for a time and under conditions sufficient to increase or decrease Csl activity.

Modulation of said activity by the administration of an agent to a mammal can be achieved by one of several techniques, including but in no way limited to introducing into said mammal a proteinaceous or non-proteinaceous molecule which:

(i) modulates expression of *Csl*;

(ii) functions as an antagonist of Csl;

(iii) functions as an agonist of Csl.

Said proteinaceous molecule may be derived from natural or recombinant sources including fusion proteins or following, for example, natural product screening. Said non-

proteinaceous molecule may be, for example, a nucleic acid molecule or may be derived from natural sources, such as for example natural product screening or may be chemically synthesised. The present invention contemplates chemical analogs of Csl capable of acting as agonists or antagonists of Csl. Chemical agonists may not necessarily be derived from Csl but may share certain conformational similarities. Alternatively, chemical agonists may be specifically designed to mimic certain physiochemical properties of Csl. Antagonists may be any compound capable of blocking, inhibiting or otherwise preventing Csl from carrying out their normal biological functions. Antagonists include monoclonal antibodies specific for Csl, or parts of Csl, and antisense nucleic acids which prevent transcription or translation of *Csl* genes or mRNA in mammalian cells. Agonists and antagonists of Csl should be understood to include any molecule which synergises with Csl to either upregulate or downregulate, respectively, its activity. For example, IGF-1 synergises with Csl to induce muscle cell hypertrophy.

The Csl, *Csl* or agent used may also be linked to a targeting means such as a monoclonal antibody, which provides specific delivery of the Csl, *Csl* or agent to the target cells.

In a preferred embodiment of the present invention, the Csl, *Csl* or agent used in the method is linked to an antibody specific for said target cells to enable specific delivery to these cells.

Administration of the Csl, *Csl* or agent, in the form of a pharmaceutical composition, may be performed by any convenient means. Csl, *Csl* or agent of the pharmaceutical composition are contemplated to exhibit therapeutic activity when administered in an amount which depends on the particular case. The variation depends, for example, on the human or animal and the Csl, *Csl* or agent chosen. A broad range of doses may be applicable. Considering a patient, for example, from about 0.1 μg to about 10 mg of Csl or agent may be administered per kilogram of body weight per day. For example from about 0.1 μg-5 mg, 10 μg-5 mg or 100 μg-1 mg. Dosage regimes may be adjusted to provide the optimum therapeutic response. For example, several divided doses may be administered daily, weekly, monthly or other suitable time intervals or the dose may be proportionally reduced as indicated by the exigencies of the situation. The Csl or agent may be administered in a convenient manner such as by the oral, intravenous (where water soluble), intranasal, intraperitoneal,

intramuscular, subcutaneous, intradermal or suppository routes or implanting (e.g. using slow release molecules). With particular reference to use of Csl or agent, these peptides may be administered in the form of pharmaceutically acceptable nontoxic salts, such as acid addition salts or metal complexes, e.g. with zinc, iron or the like (which are considered as salts for purposes of this application). Illustrative of such acid addition salts are hydrochloride, hydrobromide, sulphate, phosphate, maleate, acetate, citrate, benzoate, succinate, malate, ascorbate, tartrate and the like. If the active ingredient is to be administered in tablet form, the tablet may contain a binder such as tragacanth, corn starch or gelatin; a disintegrating agent, such as alginic acid; and a lubricant, such as magnesium stearate.

A further aspect of the present invention relates to the use of the invention in relation to mammalian disease conditions. The present invention is particularly useful, but in no way limited to, use in attenuating muscle fraility, treating muscular and myotonic dystrophies the prevention of cardiomyopathy, and treating maladaptive hypertrophy.

For example, the decline in function and restriction of adaptability of skeletal muscle is a hallmark of ageing in humans, leading to severe deficit in performance and ultimately to fraility. Maintenance of muscle integrity via modulation of Csl expression either alone or in concert with hormonal stimulation provides an approach for the prevention of muscle atrophy in ageing patients.

Muscular dystrophies are characterised by progressive weakness in specific muscle groups, leading to deterioration of muscle ultrastructures. Many of these pathologies arise from genetic defects that effect muscle integrity. Modulation of Csl activity provides an approach for counteracting the deterioration of muscle structure in some dystrophic patients. The inability to regenerate muscle tissue is another serious complication in muscular dystrophy. By utilising Csl as the therapeutic molecule the integrity of newly formed muscle cells can be preserved thereby counteracting muscle degeneration. Csl therapy is particularly attractive in gene therapy programs due to its small size and suitablity to insertion into gene therapy vectors.

Yet another example of the use of the present invention relates to cardiomyopathy. Hypertrophy, whether familial or generated by hypertension or other idiopathic causes, is a normal adaptation to increased work load in the heart during which there is new myofilament protein synthesis and deposition of new myofilament proteins into the contractile apparatus. However, if the adaptive response is excessive, the heart becomes unable to cope with the changes, and both myofibrillar and myocyte disarray result, with a high risk of sudden death for the individual. Csl functions in organising myofilaments for the purpose of protecting or elevating the pathology associated with cardiac hypertrophy.

Accordingly, another aspect of the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of an agent capable of modulating the expression of *Csl* or a derivative, homologue or mimetic thereof for a time and under conditions sufficient to modulate muscle cell development.

The term "development" is used herein in its broadest sense and includes reference to all aspects of cellular development including, but not limited to, homeostasis, viability, proliferation, migration, morphology and differentiation of cells. "Differentiation" refers to the partial or complete maturation of a cell and may be evidenced, for example, by altered cell surface protein expression.

In another aspect the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of an agent capable of modulating the activity of Csl or a derivative, homologue or mimetic thereof for a time and under conditions sufficient to modulate muscle cell development.

In another aspect the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of Csl or a derivative, homologue or mimetic thereof for a time and under conditions sufficient to modulate muscle cell development.

Yet another aspect the present invention relates to a method of treating a mammal said method comprising administering to said mammal an effective amount of *Csl* or a derivative, homologue or mimetic thereof for a time and under conditions sufficient to modulate muscle cell development.

In yet another aspect the present invention relates to the use of an agent capable of modulating the expression of *Csl* or a derivative, homologue or mimetic thereof in the manufacture of a medicament for the modulation of muscle cell development.

Another aspect of the present invention relates to the use of an agent capable of modulating the expression of Csl or a derivative, homologue or mimetic thereof in the manufacture of a medicament for the modulation of muscle cell development.

A further aspect of the present invention relates to the use of Csl or *Csl* in the manufacture of a medicament for the modulation of muscle cell development.

Still yet another aspect of the present invention relates to agents for use in modulating *Csl* or a derivative, homologue or mimetic thereof expression wherein modulating expression of said *Csl* modulates muscle cell development.

A further aspect of the present invention relates to agents for use in modulating Csl or a derivative, homologue or mimetic thereof expression wherein modulating expression of said Csl modulates muscle cell development.

Another aspect of the present invention relates to Csl or *Csl* for use in modulating muscle cell development.

Preferably said muscle cell differentiation is heart or skeletal cell differentiation and even more preferably heart cell development.

In a related aspect of the present invention, the mammal undergoing treatment may be human or an animal in need of therapeutic of prophylactic treatment.

In yet another further aspect the present invention contemplates a pharamceutical composition *Csl*, Csl or an agent capable of modulating *Csl* expression or Csl activity together with one or more pharamaceutically acceptable carriers and/or diluents. *Csl*, Csl or said agent are referred to as the active ingredients.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions (where water soluble) and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating such as licithin, by the maintenance of the required particle size in the case of dispersion and by the use of superfactants. The preventions of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimersal and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredient into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred

methods of preparation are vacuum drying and the freeze-drying technique which yield a powder of the active ingredient plus any additional desired ingredient from previously sterile-filtered solution thereof.

When *Csl*, Csl and Csl modulators are suitably protected they may be orally administered, for example, with an inert diluent or with an assimilable edible carrier, or it may be enclosed in hard or soft shell gelatin capsule, or it may be compressed into tablets, or it may be incorporated directly with the food of the diet. For oral therapeutic administration, the active compound may be incorporated with excipients and used in the form of ingestible tablets, buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 1% by weight of active compound. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 5 to about 80% of the weight of the unit. The amount of active compound in such therapeutically useful compositions in such that a suitable dosage will be obtained. Preferred compositions or preparations according to the present invention are prepared so that an oral dosage unit form contains between about 0.1 µg and 2000 mg of active compound.

The tablets, troches, pills, capsules and the like may also contain the following: A binder such as gum tragacanth, acacia, corn starch or gelatin; excipients such as dicalcium phosphate; a disintegrating agent such as corn starch, potato starch, alginic acid and the like; a lubricant such as magnesium stearate; and a sweetening agent such a sucrose, lactose or saccharin may be added or a flavouring agent such as peppermint, oil of wintergreen, or cherry flavouring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar or both. A syrup or elixir may contain the active compound, sucrose as a sweetening agent, methyl and propylparabens as preservatives, a dye and flavouring such as cherry or orange flavour. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compound may

be incorporated into sustained-release preparations and formulations.

Pharmaceutically acceptable carriers and/or diluents include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, use thereof in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

It is especially advantageous to formulate parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the mammalian subjects to be treated; each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the novel dosage unit forms of the invention are dictated by and directly dependent on (a) the unique characteristics of the active material and the particular therapeutic effect to be achieved, and (b) the limitations inherent in the art of compounding such an active material for the treatment of disease in living subjects having a diseased condition in which bodily health is impaired.

The principal active ingredient is compounded for convenient and effective administration in effective amounts with a suitable pharmaceutically acceptable carrier in dosage unit form as hereinbefore disclosed. A unit dosage form can, for example, contain the principal active compound in amounts ranging from 0.5 µg to about 2000 mg. Expressed in proportions, the active compound is generally present in from about 0.5 µg to about 2000 mg/ml of carrier. In the case of compositions containing supplementary active ingredients, the dosages are determined by reference to the usual dose and manner of administration of the said ingredients.

The pharmaceutical composition may also comprise genetic molecules such as a vector capable of transfecting target cells where the vector carries a nucleic acid molecule capable of modulating *Csl* expression or Csl activity. The vector may, for example, be a viral vector.

Still another aspect of the present invention is directed to antibodies to Csl including catalytic antibodies. Such antibodies may be monoclonal or polyclonal and may be selected from naturally occurring antibodies to Csl or may be specifically raised to Csl. In the case of the latter, Csl may first need to be associated with a carrier molecule. The antibodies and/or recombinant Csl of the present invention are particularly useful as therapeutic or diagnostic agents. Alternatively, fragments of antibodies may be used such as Fab fragments. Furthermore, the present invention extends to recombinant and synthetic antibodies and to antibody hybrids. A "synthetic antibody" is considered herein to include fragments and hybrids of antibodies. The antibodies of this aspect of the present invention are particularly useful for immunotherapy and may also be used as a diagnostic tool or for monitoring the program of a therapeutic regime.

Csl can also be used to screen for naturally occurring antibodies to Csl.

Specific antibodies can be used to screen for Csl proteins. The latter would be important, for example, as a means for screening for levels of Csl in a cell extract or other biological fluid or purifying Csl made by recombinant means from culture supernatant fluid. Techniques for the assays contemplated herein are known in the art and include, for example, sandwich assays, ELISA and flow cytometry.

It is within the scope of this invention to include any second antibodies (monoclonal, polyclonal or fragments of antibodies) directed to the first mentioned antibodies discussed above. Both the first and second antibodies may be used in detection assays or a first antibody may be used with a commercially available anti-immunoglobulin antibody. An antibody as contemplated herein includes any antibody specific to any region of Csl.

Both polyclonal and monoclonal antibodies are obtainable by immunization with the protein or peptide derivatives and either type is utilizable for immunoassays. The methods of obtaining both types of sera are well known in the art. Polyclonal sera are less preferred but are relatively easily prepared by injection of a suitable laboratory animal with an effective amount of Csl, or antigenic parts thereof, collecting serum from the animal, and isolating specific sera by any of the known immunoadsorbent techniques. Although antibodies produced by this method are utilizable in virtually any type of immunoassay, they are generally less favoured because of the potential heterogeneity of the product.

The use of monoclonal antibodies in an immunoassay is particularly preferred because of the ability to produce them in large quantities and the homogeneity of the product. The preparation of hybridoma cell lines for monoclonal antibody production derived by fusing an immortal cell line and lymphocytes sensitized against the immunogenic preparation can be done by techniques which are well known to those who are skilled in the art. (See, for example Douillard and Hoffman, Basic Facts about Hybridomas, in *Compendium of Immunology* Vol II, ed. by Schwartz, 1981; Kohler and Milstein, *Nature 256*: 495-499, 1975; *European Journal of Immunology 6*: 511-519, 1976).

Another aspect of the present invention contemplates a method for detecting Csl in a biological sample from a subject said method comprising contacting said biological sample with an antibody specific for Csl or its derivatives or homologs for a time and under conditions sufficient for an antibody-Csl complex to form, and then detecting said complex.

The presence of Csl may be determined in a number of ways such as by Western blotting, ELISA or flow cytometry procedures. These, of course, include both single-site and two-site or "sandwich" assays of the non-competitive types, as well as in the traditional competitive binding assays. These assays also include direct binding of a labelled antibody to a target.

Sandwich assays are among the most useful and commonly used assays and are favoured for use in the present invention. A number of variations of the sandwich assay technique exist,

and all are intended to be encompassed by the present invention. Briefly, in a typical forward assay, an unlabelled antibody is immobilized on a solid substrate and the sample to be tested brought into contact with the bound molecule. After a suitable period of incubation, for a period of time sufficient to allow formation of an antibody-antigen complex, a second antibody specific to the antigen, labelled with a reporter molecule capable of producing a detectable signal is then added and incubated, allowing time sufficient for the formation of another complex of antibody-antigen-labelled antibody. Any unreacted material is washed away, and the presence of the antigen is determined by observation of a signal produced by the reporter molecule. The results may either be qualitative, by simple observation of the visible signal, or may be quantitated by comparing with a control sample containing known amounts of hapten. Variations on the forward assay include a simultaneous assay, in which both sample and labelled antibody are added simultaneously to the bound antibody. These techniques are well known to those skilled in the art, including any minor variations as will be readily apparent. In accordance with the present invention the sample is one which might contain Csl including cell extract, tissue biopsy or possibly serum, saliva, mucosal secretions, lymph, tissue fluid and respiratory fluid. The sample is, therefore, generally a biological sample comprising biological fluid but also extends to fermentation fluid and supernatant fluid such as from a cell culture.

In the typical forward sandwich assay, a first antibody having specificity for the Csl or antigenic parts thereof, is either covalently or passively bound to a solid surface. The solid surface is typically glass or a polymer, the most commonly used polymers being cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene. The solid supports may be in the form of tubes, beads, discs of microplates, or any other surface suitable for conducting an immunoassay. The binding processes are well-known in the art and generally consist of cross-linking covalently binding or physically adsorbing, the polymer-antibody complex is washed in preparation for the test sample. An aliquot of the sample to be tested is then added to the solid phase complex and incubated for a period of time sufficient (e.g. 2-40 minutes) and under suitable conditions (e.g. 25°C) to allow binding of any subunit present in the antibody. Following the incubation period, the antibody subunit solid phase is washed and dried and incubated with a second antibody specific for a portion of the hapten.

The second antibody is linked to a reporter molecule which is used to indicate the binding of the second antibody to the hapten.

An alternative method involves immobilizing the target molecules in the biological sample and then exposing the immobilized target to specific antibody which may or may not be labelled with a reporter molecule. Depending on the amount of target and the strength of the reporter molecule signal, a bound target may be detectable by direct labelling with the antibody. Alternatively, a second labelled antibody, specific to the first antibody is exposed to the target-first antibody complex to form a target-first antibody-second antibody tertiary complex. The complex is detected by the signal emitted by the reporter molecule.

By "reporter molecule" as used in the present specification, is meant a molecule which, by its chemical nature, provides an analytically identifiable signal which allows the detection of antigen-bound antibody. Detection may be either qualitative or quantitative. The most commonly used reporter molecules in this type of assay are either enzymes, fluorophores or radionuclide containing molecules (i.e. radioisotopes) and chemiluminescent molecules.

In the case of an enzyme immunoassay, an enzyme is conjugated to the second antibody, generally by means of glutaraldehyde or periodate. As will be readily recognized, however, a wide variety of different conjugation techniques exist, which are readily available to the skilled artisan. Commonly used enzymes include horseradish peroxidase, glucose oxidase, beta-galactosidase and alkaline phosphatase, amongst others. The substrates to be used with the specific enzymes are generally chosen for the production, upon hydrolysis by the corresponding enzyme, of a detectable color change. Examples of suitable enzymes include alkaline phosphatase and peroxidase. It is also possible to employ fluorogenic substrates, which yield a fluorescent product rather than the chromogenic substrates noted above. In all cases, the enzyme-labelled antibody is added to the first antibody hapten complex, allowed to bind, and then the excess reagent is washed away. A solution containing the appropriate substrate is then added to the complex of antibody-antigen-antibody. The substrate will react with the enzyme linked to the second antibody, giving a qualitative visual signal, which may be further quantitated, usually spectrophotometrically, to give an

indication of the amount of hapten which was present in the sample. "Reporter molecule" also extends to use of cell agglutination or inhibition of agglutination such as red blood cells on latex beads, and the like.

Alternately, fluorescent compounds, such as fluorecein and rhodamine, may be chemically coupled to antibodies without altering their binding capacity. When activated by illumination with light of a particular wavelength, the fluorochrome-labelled antibody adsorbs the light energy, inducing a state to excitability in the molecule, followed by emission of the light at a characteristic color visually detectable with a light microscope. As in the EIA, the fluorescent labelled antibody is allowed to bind to the first antibody-hapten complex. After washing off the unbound reagent, the remaining tertiary complex is then exposed to the light of the appropriate wavelength the fluorescence observed indicates the presence of the hapten of interest. Immunofluorescene and EIA techniques are both very well established in the art and are particularly preferred for the present method. However, other reporter molecules, such as radioisotope, chemiluminescent or bioluminescent molecules, may also be employed.

The present invention also contemplates genetic assays such as involving PCR analysis to detect *Csl* or its derivatives.

Further features of the present invention are more fully described in the following examples. It is to be understood, however, that this detailed description is included solely for the purposes of exemplifying the present invention. It should not be understood in any way as a restriction on the broad description of the invention as set out above.

SUMMARY OF SEQ ID Nos.

Sequence	SEQ ID NO.
cDNA nucleotide sequence of murine Csl	1
amino acid sequence of murine Csl	2
cDNA nucleotide sequence of human Csl	3
amino acid sequence of human Csl	4
amino acid sequence of Xenopus Csl	5

EXAMPLE 1

CLONING AND SEQUENCE ANALYSIS

In order to isolate novel genes that may be regulated by Nkx2.5 in the heart a modified form of representational difference analysis (RDA) 'PCR-select' (Clontech Ltd) was used. Mice heterozygous for the mutant allele Nkx2.5-/+ were intercrossed and pregnant mothers were killed at 8.5 days post-coitum (dpc). Heart tissue was isolated from each fetus and used to extract RNA. The remaining body tissue was used to isolate DNA for PCT genotyping. Differential hybridisations were performed according to the manufacturers instructions using cDNA synthesised from Nkx2.5-/- heart RNA as the driver and heart RNA wild-type for Nkx2.5 as the tester. The resulting PCR bands were isolated, cloned, sequenced and hybridised against the original cDNA isolates to determine differential expression. It was established that one of these products (Cs1) is found in normal fetal heart cDNA but is either very low or absent from Nkx2.5/- heart cDNA. The partial cDNA clone was compared against sequence databases and a series of expressed sequence tags (ESTs) were found to match at highly significant levels. These are listed in Table 2. These ESTs were derived from random sequencing projects of fetal mouse heart and skeletal muscle and human heart and skeletal muscle cDNA libraries. By compiling contigs of ESTs it was possible to predict a full length spliced product for mouse and human genes. Primers were designed to the termini of the predicted mouse cDNA and a product was amplified by PCR of the correct size from fetal heart cDNA. This was cloned, sequenced and confirmed as identical to the EST contig. The open reading frame (ORF) of the mouse sequence encodes a short protein of 86 amino acids which has no significant homology to any known protein in the current databases. The small ORF was compared with the predicted human cDNA. Due to slight sequence differences, the human cDNA has only one ORF of reasonable length that could encode a similar protein to the mouse. The human sequence encodes an 87 amino acid polypeptide that is 86% identical to the mouse (Figure 3). The human gene contains 5 exons covering 50 kb and the splice sites, splice orientation and exon sequence matches have been predicted from the EST cDNA configuration.

The mouse cDNA clone was used to isolate several cDNA clones from a Xenopus cDNA library.

EXAMPLE 2

EXPRESSION ANALYSIS

The distribution of Cs1 mRNA during early mouse development was analysed in detail using whole mount *in-situ* hybridisation with an RNA probe synthesised from the full length cDNA. Message was detected in the cardiac crescent at 8.0 dpc which persists into the linear heart tube stage. During heart looping, the expression is seen exclusively in the myocardial layer of the outer curvature of the presumptive left and right ventricles. Expression can be seen later in discrete regions of the developing atria but remained off in atrioventricular canal. Cs1 is a novel marker for these regions since this expression pattern has not been previously reported. It is believed that Csl expression may define the subpopulation of cells that form the definitive contractile tissue of the ventricles and atria, rather than its associated elements such as the valves and the septum. RNase protection analysis has revealed that in the adult, Cs1 is expressed strongly in both the heart and skeletal muscle. Skeletal muscle expression does not appear to be specific to, or biased in, fast or slow muscle fibre types since Northern analysis shows relatively even mRNA levels in the seven different muscle groups tested to date. However, the level of muscle Cs1 mRNA does appear to diminish in ageing mice.

In developing limb buds, Cs1 mRNA can be detected as early as 12.5 dpc by RNase protection analysis.

An analysis has been performed of Csl expression in myogenic cell lines, which can be propogated in high serum as undifferentiated myoblasts, or induced to withdraw from the cell cycle and differentiate by removal of serum. In the robust and commonly used C2C12 line, thought to represent adult satalite cells, *Csl* was strongly expressed only after cell cycle withdrawal. Thus, *Csl* appears to function in the differentiation phase of muscle development. *Csl* was not expressed in a compromised myogenic cell line, L6E9, derived

from neonatal limb muscle. L6E9 cells differentiate very poorly and only fuse into myotubes after long periods in culture. However, in the presence of IGF1, which is normally expressed in the microenvironment of developing muscles, differentiation and fusion is dramatically improved. In the presence of IGF1, Csl is activated, although very weakly compared to levels in C2C12 cells.

Csl expression is also analysed by *in-situ* hybridisation to sectioned material. Firstly, sections of later stage mouse fetuses are examined for the changing pattern of *Csl* distribution in the developing heart to determine, in particular, the progression of the atrial *Csl* mRNA deposition. Secondly, the origin of expression found in the adult heart is determined from sections. Finally, a time course of sections from late-stage embryos are examined for the onset of expression in axial skeletal muscle and limb buds and compared with results derived from ongoing RNase protection analyses.

EXAMPLE 3

TARGETED DISRUPTION OF THE CS1 ALLELE

The cosmid clones containing the genomic sequences of human CSL were derived from the long arm of the human X chromosome. X-unique genes are known to be highly conserved on the X chromosomes of all mammals due to the constraints of dosage compensation (7). Mouse Csl have been mapped to the mouse X by comparing the intensity of bands hybridising to a Csl probe on Southern blots of equally loaded male and female genomic DNA. As expected, the intensity of the bands in the female lanes is twice that of the male lanes.

EXAMPLE 4

CREATION OF A NULL MUTATION OF *Csl*

The X chromosomal location of the gene complicates the issue since the embryonic stem (ES) cell line used to make knockouts is XY and is therefore hemizygous for *Csl*. If a null mutation is introduced into the single copy of *Csl*, all *Csl* expression is lost. Assuming that

lack of *Csl* expression is significantly deleterious, chimeras made from the targeted cell line mixed with wild-type blastocysts may not survive if the proportion of the targeted cells is very high as it needs to be to achieve germ line transmission. Under these circumstances the effect of the knockout is studied in fetal chimeras.

A conditional knockout construct has been designed that is intended to maintain unaltered *Csl* expression in the primary targeted conformation. The introduction of loxP sites from the coliphage P1 system (8) allows the later excision of a portion of this construct on exposure to Cre recombinase. The construct is built to replace the second exon which contains a short stretch of 5' untranslated region, the initiating methionine codon and the amino terminus of the predicted protein which by comparison with human and Xenopus sequences appears to be the most conserved domain. The construct does not remove any exon or intron sequence but introduces one loxP site within the 5' untranslated region and another within the intron between exon 2 and 3. A PGK-neo selection cassette has been introduced into the intron between the two loxP sites and a human placental alkaline phosphatase cDNA has been placed after the intronic loxP site. Cre recombination excises the intervening sequence between the two loxP sites thus removing the selection cassette, the 5' portion of *Csl* coding sequence and the splice donor of exon 2. The coding sequence of alkaline phosphatase is brought into the spliced transcript in a position previously occupied by the initiation codon of *Csl* and if a stable mRNA is formed which is translated it allows the histochemical detection of alkaline phosphatase in cells that have undergone Cre excision. In any event, *Csl* translation is destroyed. Therefore, these mice are bred to females to produce Csl-targeted (*Csl*^{t}) hemizygous males and *Csl*^{t} heterozygous females which are used to maintain the line. To produce null mutants, male hemizygous *Csl*^{t} mice or the founder chimeras are mated to transgenic females expressing Cre in the germ line (9). Since this Cre transgene is inserted on the X chromosome it is necessary to breed a second generation from the femal progeny of this cross because only the females carry both the *Csl*^{t} allele and the Cre transgene. Males produced from this second generation all carry the null mutation whereas the females are heterozygous null. Alternatively, the *Csl* knockout is activated in a tissue-specific fashion by breeding *Csl*^{t} mice with lines that express the Cre protein under the control of promoters that are active in the embryonic heart or skeletal

muscle.

EXAMPLE 5

ANALYSIS OF PROTEIN LOCALISATION, STRUCTURE AND BIOCHEMISTRY

The Csl protein appears to be tethered to a component of the cytoskeleton, with the pattern in fibroblasts resembling that of the microfilament or intermediate filament networks. To produce a form of chisel protein detectable in cells the cDNA was cloned as a fusion with the flag epitope in pEFBOS. This construct was transiently transfected into cos cells which were fixed, permeablised and stained using anti-flag antibodies. In these experiements, staining was found in both the nuclei and the cytoplasm although in many cells staining appeared to be excluded from the nucleus. Some cells appeared to contain filamentous strands of flourescence that could reflect association with cytoskeletal elements. To pursue this result, stable 10T1/2 transfectants were made using the same construct and 24 independent clones were isolated. In these cell lines, Csl appeared to be present in high density in and around the nucleus and in a gradually diminishing concentration out to the periphery of the cell. In the distal regions of the cytoplasm where the staining was weakest, flourescence appeared to be filamentous and was often stronger near sites of cytoplasmic projections.

The stable 10T1/2 cell lines are stained with anti-flag antibodies after exposure of the cells to agents that disrupt specific elements of the cytoskeleton including cytocholasin B which disrupts actin microfilaments and colcemid which disrupts microtubules. Secondly the cells are co-stained with a series of commercially available antibodies raised against actin associated proteins, tubulin associated proteins and motor proteins such as myosin. Both of these sets of experiments are performed with and without triton extraction prior to fixation and staining which may reveal underlying structure that is obscured by an overabundance of flag-tagged *Csl* protein. These cells are easily induced to fuse and form myotubes in culture whereupon the deposition of flag-*Csl* is visualised.

Polyclonal antisera are raised against whole Csl protein produced by fusion with glutathione-s-transferase (GST) in a pGEX vector (Pharmacia Ltd) followed by affinity-purification and thrombin cleavage. *Csl* mRNA is detectable during the differentiation of C2C12 cells from myoblasts into fused myotubes (see functional analyses *in vitro*) therefore these cells are used to examine the endogenous localisation of chisel protein.

The 10T1/2 and C2C12 Flag-*Csl* stable cell lines are also used in coimmunoprecipitation assays to determine if *Csl* binds to know proteins. Sepharose bound flag antibody are used to precipitate the Flag-*Csl* fusion protein from cell lysates and the resulting products electrophoresed. The coimmunoprecipitates are visualised initially by ^{35}S methionine labelling. A wide range of antibodies directed against potential binding partners are tested on western blots of these coimmunoprecipitates.

The amino acid sequence of chisel has no significant homology to any known proteins in the database. However, its small size and the absence of cysteine residues makes it a good candiate for analysis by nuclear magnetic resonance (NMR) to solve the 3 dimensional structure. The GST-*Csl* fusion construct is used in a large scale protein production for NMR spectroscopy.

Analysis of the mouse Csl protein sequence using a structure prediction algorithm indicates that Csl has homology with the essential myosin light chain of scallop muscle. This homology lies over that part of the light chain which bind to myosin heavy chain in the actomyosin complex suggesting that Csl has myosin-binding capacity.

Some idea of the relative importance of domains within the chisel protein can be obtained by interspecific comparisons and the alignment of mouse and human with Xenopus sequences. Comparisons between mouse, human and frog Csl suggest a tripartite organisation, with the N-terminal region being more conserved than C-terminal and central regions. A series of mutations across the Csl protein are made and their characteristics in cell culture assays assessed. Their localization within the cell (cytoskeletal tethering; nuclear/cytoplasmic localisation; co-localization with myosin), their ability to enhance differentiation and fusion

in C2C12 and L6E9-MLC3-IGF cells in low serum, and their ability to induce precocious differentiation in L6E9-MLC3-IGF cell in high serum are assessed. Mutant protection are also assessed for their ability to bind to individual cellular or cytoskeletal components, known and unknown, *in vitro*. Site-directed mutagenesis is performed using the QuickChange system (Stratagene Ltd). Mutations are introduced into the FLAG-Csl and GST-Csl expression vectors and transfected into cell lines, or used for *in vitro* binding assays, respectively.

EXAMPLE 6

IN-VITRO ASSAYS OF *Csl* FUNCTION

The *Csl* targeted ES cell line that is used to create null mutant mice is also differentiated in culture to produce embryoid bodies. Under these conditions it is possible to induce considerable levels of myogenic differentiation. By exposure to Cre through transfection of a Cre construct the single targeted *Csl* locus is inactivated and the effect of *Csl* loss on myogenic differentiation observed by testing for levels of specific markers and the effect on cell phenotype. The pressure of ore-excised cells is assayed by alkaline phosphatase production.

Cell line models of cardiac myocytes are widely regarded as unrepresentative and some are known to lack common cardiac specific markers such as *Nkx*2.5. Primary cultures of mouse cardiac myocytes extracted from 10-12 dpc fetal heart tissue are produced and studied. The effect of *Csl* overexpression is studied in these cells by transfection with expression constructs. Initially, morphological effects and cell proliferation rates are assessed, followed by an examination of specific cardiac marker gene alterations.

The C2C12 cell line was established from mouse adult leg muscles which were crush-injured to induce satellite cell proliferation (11). These cells maintain a fibroblast-like phenotype in media containing high concentrations of fetal calf serum, but rapidly withdraw from the cell cycle and undergo myogenic differentiation when the media is replaced with lower concentrations of horse serum. These cells express a wide variety of adult muscle-specific

markers including the myogenic factors and insulin-like growth factors. In C2C12 cells, which normally expresses Csl at high levels when induced to differentiate, enforced expression of FLAG-Csl had no significant impact on cell morphology during proliferation in high serum, although a minor enhancement of cell fusion and/or hypertrophy after serum withdrawal was seen. However, Western analysis showed that myosin heavy chain levels were increased approximately 10 fold during differentiation, relative to a tubulin control.

In L6E9 cells, which express low levels of Csl in the presence of IGF1, the outcome was much more dramatic. Enforced expression of FLAG-Csl in L6E9 cells, which also carry a stably integrated transgene expressing IGF1 from a muscle differentiation-specific promoter (the myosin light chain 2 promoter; MLC3), had two effects. First, in these cells (L6E9-MLC3-IGF) Csl induced a massive hypertrophy, as judged by the dramatic increase in the size of myotubes at early differentiation time points. We estimate that Csl-expressing cells form myotubes ("megatubes") that are at least 10 fold the diameter of those formed from L6E9-MLC3-IGF cells alone. Second, even in the presence of high serum, which confers multiple inhibitory signals for myogenic differentiation, FLAG-Csl-expressing L6E9-MLC3-IGF cells formed myosin-positive multinucleate myotubes at a significant frequency, while L6E9-MLC3-IGF cells also could not differentiation. At present, the frequency of this effect is hard to quantitiate, because even under constant positive drug selection for the Csl vector, cultures became completely devoid of Csl-expressing cells within 4 passages. We interpret this to mean that the Csl-expressing cells are lost rapidly from the culture due to differentiation. Thus, Csl can in some way override the inhibitory signals for differentiation conferred by high serum and activate the myogenic program. In L6E9-MLC3-IGF cells expressing FLAG-Csl, myosin heavy chains and the FLAG epitope co-localise, further evidence that Csl can bind myosins.

An antisense *Csl* construct in a eukaryotic expression vector is made which is stably integrated into C2C12 cells. Further, an inducible *Csl* expression construct is made because the premature withdrawal from proliferation seen in cotransfected MLC1-IGF and Flag-*Csl* cells has prevented the creation of stable cell lines overexpressing both genes. The *Csl*-induced hypertrophic response in MLC1-IGF L6 cells is also used as a functional bioassay

to test, for example, the effectiveness of mutated *Csl* proteins.

Transgenic *Csl* mice made with an alpha actin promoter are produced so that the effect of *Csl* overexpression *in-vivo* is observable, particularly with regard to muscle hypertrophy and muscle ageing. Also cell synchronisation studies are used to examine whether the exclusion of *Csl* from the nucleus observed in earlier transfection experiments is related to cell cycle.

EXAMPLE 7

MYOGENIC ACTIVITY OF Csl IN CELL CULTURE ASSAYS

A full molecular workup is performed on the C2C12 and L6E9-MLC3-IGF cells in which Csl has been over-expressed. Changes in expression of a range of myofilament genes (thick versus thin filament; embryonic versus adult isoforms), are examined as well as those encoding adhesion molecules, cytoskeletal proteins and myogenic regulators of the myod and Mef2 families. To avoid loss of differentiating cells during passage in cell culture, the cell lines with Csl expressed from an inducible promoter and/or a differentiation-specific promoter (MLC3) are rederived. An antisense Csl vector is constructed to examine the effects of reduce Csl levels on differentiation in culture. The fact that L6E9 and L6E9-MLC3-IGF cells express Csl poorly or not at all, and that over-expressed Csl enhances differentiation in L6E9-MLC3-IGF cells, suggests it has an essential function in myogenesis. Cells expressing antisense constructs are monitored for Csl expression and are characterised macroscopically and with the full range of molecular markers.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

TABLE 2:

W38301	: zc15c11.r1 Soares parathyroid ...	+1	415	8.3e-55	2	H
AA033164	: mi37c11.r1 Soares mouse embryo...	+1	440	1.2e-53	1	M
W18646	: mb98a05.r1 Soares mouse p3NMF1...	+2	379	1.2e-53	2	M
AA434782	: ve23c01.r1 Soares mouse NbMH M...	+2	227	7.7e-53	2	M
AA060214	: mj65h06.r1 Soares mouse p3NMF1...	+2	432	1.4e-52	1	M
W18392	: mb88e01.r1 Soares mouse p3NMF1...	+1	424	1.8e-51	1	M
W36988	: mb65b11.r1 Soares mouse p3NMF1...	+2	424	1.9e-51	1	M
W29828	: mc07b11.r1 Soares mouse p3NMF1...	+2	313	5.6e-44	2	M
W13738	: mb32a12.r1 Soares mouse p3NMF1...	+1	227	9.3e-43	2	M
AA211521	: zn55b01.r1 Stratagene muscle 9...	+1	183	5.5e-36	2	H
AA389647	: M104 Fetal heart, Lambda ZAP E...	+2	159	6.1e-35	2	H
AA214155	: zn58f10.r1 Stratagene muscle 9...	+3	183	5.6e-34	3	H
W97451	: mf97c02.r1 Soares mouse embryo...	+1	286	1.0e-32	1	M
W29186	: mc22f02.r1 Soares mouse p3NMF1...	+3	241	1.2e-26	1	M
W07478	: za96c10.r1 Soares fetal lung N...	+2	220	9.4e-24	1	H
R58129	: F9260 Fetal heart Homo sapiens...	+1	206	3.1e-21	1	H
AA094015	: c11599.seq.F Fetal heart, Lamb...	+3	157	6.4e-21	2	H
W14689	: mb34a08.r1 Soares mouse p3NMF1...	+3	197	1.5e-20	1	M
HUMU228D4	: Human cosmid U228D4, complete ...	-2	147	7.0e-17	3	H Gen
N87511	: LL1191F Fetal heart, Lambda ZA...	+1	169	7.2e-17	1	H
HUMU112E8	: Human cosmid U112E8, complete ...	-1	179	3.4e-16	1	H Gen
AA248485	: csh0287.seq.F Human fetal hear...	+1	110	4.4e-13	2	H
N84085	: KK5613F Homo sapiens cDNA clon...	+3	100	2.5e-07	1	H
AA249723	: kk5613.seq.F Human fetal heart...	+3	100	2.5e-07	1	H
AA211443	: zn55b01.s1 Stratagene muscle 9...	-3	89	1.1e-06	2	H
N56276	: JJ5022F Homo sapiens cDNA clon...	+1	97	1.8e-06	1	H
AA249531	: jj5022.seq.F Human fetal heart...	+1	97	1.8e-06	1	H
AA214147	: zn58f02.r1 Stratagene muscle 9...	+1	97	6.3e-05	1	H
AA092554	: 115591.seq.F Fetal heart, Lamb...	+1	89	7.0e-05	1	H
AA247862	: j3408.seq.F Human fetal heart,...	+3	83	0.00078	1	H
AA249722	: kk5547.seq.F Human fetal heart...	+3	79	0.0037	1	H

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: THE WALTER AND ELIZA HALL INSTITUTE OF MEDICAL RESEARCH

(ii) TITLE OF INVENTION: NOVEL THERAPEUTIC MOLECULES AND USES THEREFOR

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: DAVIES COLLISON CAVE
(B) STREET: 1 LITTLE COLLINS STREET
(C) CITY: MELBOURNE
(D) STATE: VICTORIA
(E) COUNTRY: AUSTRALIA
(F) ZIP: 3000

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: AU PROVISIONAL
(B) FILING DATE: 27-MAR-1998
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: HUGHES, DR E JOHN L
(C) REFERENCE/DOCKET NUMBER: EJH/TDO/DK

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: +61 3 9254 2777
(B) TELEFAX: +61 3 9254 2770
(C) TELEX: AA 31787

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 932 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 199..453

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GCTCTCAGGA CTGGAGAGAG ACAGAGCACT CCAGCTATTT CAGCCACATG AAAAGCACTG      60

GAATTGAGAT CCCCGCTCAG AGGACACCGG GAGTTCCTTC TATCCTGTAA AGCGCTTTTT     120

GTGTTTTTGC ACCTGGCCGC CTGGGACTGT CCTCAGGCAG TAAACCAATC CAGAGAGCAG     180

GGCTAAGACC TTGTGAAT ATG TCG AAG CAG CCA ATT TCC AAC GTC AGA GCC        231
                    Met Ser Lys Gln Pro Ile Ser Asn Val Arg Ala
                      1               5                  10

ATC CAG GCG AAT ATC AAT ATT CCA ATG GGA GCC TTT CGT CCG GGA GCT       279
Ile Gln Ala Asn Ile Asn Ile Pro Met Gly Ala Phe Arg Pro Gly Ala
            15                  20                  25

GGG CAG CCT CCC AGA AGG AAA GAG AGT ACT CCT GAA ACT GAG GAG GGA       327
Gly Gln Pro Pro Arg Arg Lys Glu Ser Thr Pro Glu Thr Glu Glu Gly
        30                  35                  40

GCT CCT ACC ACC TCA GAG GAA AAG AAG CCA ATT CCT GGA ATG AAG AAA       375
Ala Pro Thr Thr Ser Glu Glu Lys Lys Pro Ile Pro Gly Met Lys Lys
    45                  50                  55

TTT CCA GGA CCT GTT GTC AAC TTG TCT GAG ATC CAA AAT GTT AAA AGT       423
Phe Pro Gly Pro Val Val Asn Leu Ser Glu Ile Gln Asn Val Lys Ser
60                  65                  70                  75

GAA CTG AAA TTT GTC CCC AAA GGT GAA CAG TAGTCGAAAG GACACAAAAG        473
Glu Leu Lys Phe Val Pro Lys Gly Glu Gln
                80                  85

TTCACATTGG ATGCTTAGAA TCAGGAGATG CATTTCGTTG ACGTGTTTTT CCAAGGGAGA     533

AAAAACAATG GGTTGAAATA AACAACTTCC TGAACATTTT ATACATTTGT ATGATGATCA     593

CAAACCTCCT GAATGCCCAA GACTCTAGCA AAAATATCCT GTTTGTACAT TTATATTTCT     653

TCCTTTTACT TGGTTGCATT TCTCACTTTA GCTACATTTT TGGCACCTTG TAGAGCAAAT     713

CAGCACACGA ATTTACAACC TGGGAAGTGT GGTTTTGAGG AGAGATGTGA TTTTTATGAA     773

GGGGGGGATG GCAACGTGCA AGCAGTGATT TTGATGTTAA GTACTTTAAG TTACTTCCCA     833

CGGTCTTTTG GTCAATATTT GAAATGGTTT CTTCACCTTT TAAATTATCT CAATTAACTT     893

TTTATGAGTT CAAATANATA TTTGAGTAAA TGTAAAAAN                            932
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 85 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ser Lys Gln Pro Ile Ser Asn Val Arg Ala Ile Gln Ala Asn Ile
 1               5                  10                  15

Asn Ile Pro Met Gly Ala Phe Arg Pro Gly Ala Gly Gln Pro Pro Arg
            20                  25                  30

Arg Lys Glu Ser Thr Pro Glu Thr Glu Glu Gly Ala Pro Thr Thr Ser
        35                  40                  45

Glu Glu Lys Lys Pro Ile Pro Gly Met Lys Lys Phe Pro Gly Pro Val
    50                  55                  60

Val Asn Leu Ser Glu Ile Gln Asn Val Lys Ser Glu Leu Lys Phe Val
65                  70                  75                  80

Pro Lys Gly Glu Gln
                85
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 887 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 185..448

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GGTTCTCAAT ACCGGGAGAG GCACAGAGCT ATTTCAGCCA CATGAAAAGC ATCGGAATTG      60

AGATCGCAGC TCAGAGGACA CCGGGCGCCC CTTCCACCTT CCAAGGAGCT TTGTATTCTT     120

GCATCTGGCT GCCTGGGACT TCCCTTAGGC AGTAAACAAA TACATAAAGC AGGGATAAGA     180

CTGC ATG AAT ATG TCG AAA CAG CCA GTT TCC AAT GTT AGA GCC ATC CAG       229
     Met Asn Met Ser Lys Gln Pro Val Ser Asn Val Arg Ala Ile Gln
       1               5                  10                  15

GCA AAT ATC AAT ATT CCA ATG GGA GCC TTT CGG CCA GGA GCA GGT CAA       277
Ala Asn Ile Asn Ile Pro Met Gly Ala Phe Arg Pro Gly Ala Gly Gln
                20                  25                  30

CCC CCC AGA AGA AAA GAA TGT ACT CCT GAA GTG GAG GAG GGT GTT CCT       325
Pro Pro Arg Arg Lys Glu Cys Thr Pro Glu Val Glu Glu Gly Val Pro
            35                  40                  45
```

```
CCC ACC TCG GAT GAG GAG AAG AAG CCA ATT CCA GGA GCG AAG AAA CTT       373
Pro Thr Ser Asp Glu Glu Lys Lys Pro Ile Pro Gly Ala Lys Lys Leu
         50                  55                  60

CCA GGA CCT GCA GTC AAT CTA TCG GAA ATC CAG AAT ATT AAA AGT GAA       421
Pro Gly Pro Ala Val Asn Leu Ser Glu Ile Gln Asn Ile Lys Ser Glu
     65                  70                  75

CTA AAA TAT GTC CCC AAA GCT GAA CAG TAGTAGGAAG AAAAAAGGAT             468
Leu Lys Tyr Val Pro Lys Ala Glu Gln
 80                  85

TGATGTGAAG AAATAAAGAG GCAGAAGATG GATTCAATAG CTCACTAAAA TTTTATATAT     528

TTGTATGATG ATTGTGAACC TCCTGAATGC CTGAGACTCT AGCAGAAATG GCCTGTTTGT     588

ACATTTATAT CTCTTCCTTC TAGTTGGCTG TATTTCTTAC TTTATCTTCA TTTTTGGCAC     648

CTCACAGAAC AAATTAGCCC ATAAATTCAA CACCTGGAGG GTGTGGTTTT GAGGAGGGAT     708

ATGATTTTAT GGAGAATGAT ATGGCAATGT GCCTAACGAT TTTGATGAAA AGTTTCCCAA     768

GCTACTTCCT ACAGTATTTT GGTCAATATT TGGAATGCGT TTTAGTTCTT CACCTTTTAA     828

ATTATGTCAC TAAACTTTGT ATGAGTTCAA ATAAATATTT GACTAAATGT AAAATGTGA      887
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 88 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Asn Met Ser Lys Gln Pro Val Ser Asn Val Arg Ala Ile Gln Ala
 1               5                  10                  15

Asn Ile Asn Ile Pro Met Gly Ala Phe Arg Pro Gly Ala Gly Gln Pro
            20                  25                  30

Pro Arg Arg Lys Glu Cys Thr Pro Glu Val Glu Glu Gly Val Pro Pro
        35                  40                  45

Thr Ser Asp Glu Glu Lys Lys Pro Ile Pro Gly Ala Lys Lys Leu Pro
    50                  55                  60

Gly Pro Ala Val Asn Leu Ser Glu Ile Gln Asn Ile Lys Ser Glu Leu
65                  70                  75                  80

Lys Tyr Val Pro Lys Ala Glu Gln
                85
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Met Ser Lys Gln Pro Ala Ser Asn Ile Arg Ser Ile Gln Ala Asn Ile
1               5                  10                  15

Asn Ile Pro Met Gly Ala Phe Arg Pro Gly Ala Gly Gln Pro Pro Lys
            20                  25                  30

Arg Lys Glu Phe Ser Thr Glu Glu Glu Gln His Val Pro Thr Pro Glu
        35                  40                  45

Ser Glu Glu Lys Ser Glu Glu Lys Lys Pro Ile Pro Gly Ala Val Lys
    50                  55                  60

Leu Pro Gly Pro Ala Phe Asn Leu Ser Glu Thr
65                  70                  75
```

DATED this 27th day of March, 1998

The Walter and Eliza Hall Institute
of Medical Research
by its Patent Attorneys
DAVIES COLLISON CAVE

BIBLIOGRAPHY:

1. Molkentin JD & Olson EN (1996). Curr. Opin. *Genet. Dev. 6:*445-453

2. Bodmer R. (1993). *Development 118:*719-729.

3. Lints TJ, Parsons L, Hartley L, Lyons I & Harvey RP. (1993). *Development 119:*419-431.

4. Lyons Y, Parsons LM, Hartley L, Li R, Andrews JE, Robb L & Harvey RP. (1995). *Genes Dev. 9:*1654-1666.

5. Durocher D, Chen C-Y, Ardati A, Schwartz RJ, & Nemer M. (1996). *Mol. Cell. Biol. 16:*4648-4655.

6. Zou Y, Evans S, Chen J, Kuo JC, Harvey RP & Chien KR. (1997). *Development 124:*793-804.

7. Ohno S. (1967). *"Sex chromosomes and sex-linked genes".* Springer-Verlag, Berlin

8. Sauer B, Whealy M, Robbins A & Enquist L. (1987). *Proc. Natl. Acad. Sci. USA. 84:*9108-9112.

9. Schwenk F, Baron U, Rajewsky K. (1995). *Nucleic Acids Res. 23:*5080-5081.

10. Huang D pEfbos vector.

11. Yaffe D, Saxel O. (1977). *Nature 270:*725-727.

12. Yaffe D. (1968). *Proc. Natl. Acad. Sci. USA. 61:*477-483.

13. Engert JC, Berglund EB, Rosenthal N. (1996). *J. Cell Biol. 135:*431-440.

CHISEL MOUSE GENOMIC MAP.

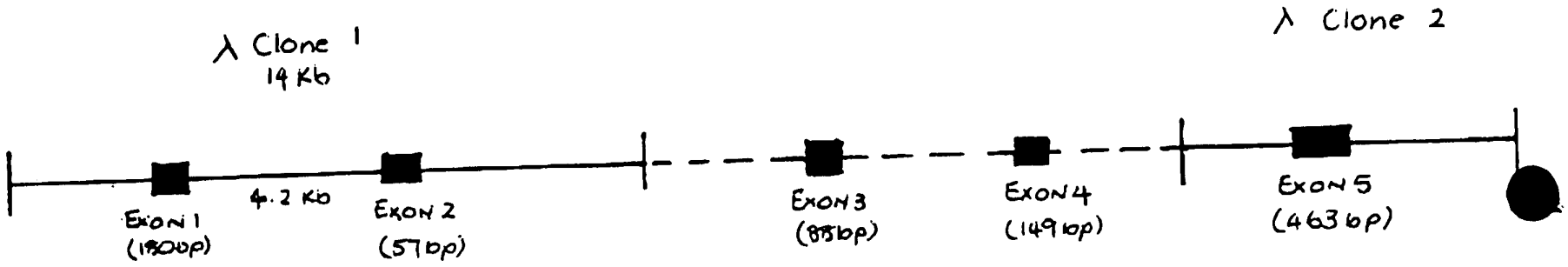

Figure 1

HISEL HUMAN GENOMIC MAP

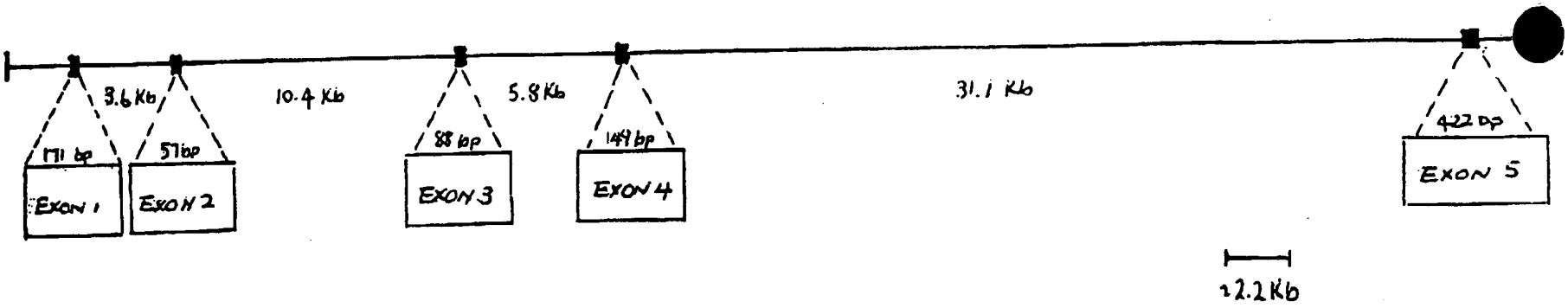

Figure 2

```
Human Csl  M S K Q P V S N V R A I Q A N I N I P M G A F R P G A G Q P P
Mouse Csl  M S K Q P I S N V R A I Q A N I N I P M G A F R P G A G Q P P
Xenopus Csl M S K Q P A S N I R S I Q A N I N I P M G A F R P G A G Q P P

Human Csl  R R K E  C T P E V E E G V P P T - S D - - - E E K K P I P G
Mouse Csl  R R K E  S T P E T E E G A P T T - S - - - - E E K K P I P G
Xenopus Csl K R K E F S T E E - E Q H V P T P E S E E K S E E K K P I P G

Human Csl  A K K L P G P A V N L S E I Q N I K S E L K Y V P K A E Q .
Mouse Csl  M K K F P G P V V N L S E I Q N V K S E L K F V P K G E Q .
Xenopus Csl A V K L P G P A F N L S E T .
```

FIGURE 3